

REMARKS

These amendments and remarks are responsive to the non-final Office Action issued on May 16, 2007. By this Response, claims 53, 57, 65 and 71 are amended, and claim 55 is cancelled without prejudice. No new matter is added. Claims 1-52, 60-64 and 66-70 were previously cancelled. Claims 53, 54, 56-59, 65 and 71 are active for examination.

The Office Action rejected claim 71 under 35 U.S.C. §112, second paragraph, as being indefinite. Claims 53-59, 65 and 71 stood rejected under 35 U.S.C. §101 for being directed to non-statutory subject matter. The Office Action rejected claims 53-59, 65 and 71 under 35 U.S.C. §103(a) as being unpatentable over Liddy (5,963,940) in view of Wermter (“Recurrent Neural Network Learning for Text Routing”). The specification and claims 53, 57, 65 and 71 were objected to.

It is submitted that the Rejections are respectfully overcome and the objection is addressed in view of the amendments and/or remarks presented herein.

The Telephone Interview

The Examiner is thanked for the courtesy for granting a telephone interview on September 11, 2007 to discuss the outstanding Office Action. The Examiner suggest overcoming the rejection under 35 U.S.C. §101 by reciting that a calculation result is displayed to create a tie to the real world. By this Response, the claims are amended as suggested by the Examiner.

The Rejections of Claim 55 Are Overcome

By this Response, claim 55 is cancelled without prejudice. Accordingly, the rejections of claim 55 are now moot.

The Objections to the Specification and Claims 53, 57, 65 and 71 Are Overcome

The Office Action objected to claims 53, 57, 65 and 71 for minor formality. Specifically, the Examiner suggested that the term “particular” should be deleted from claims 53, 37, 65 and 71. By this Response, the identified claims are amended as suggested by the Examiner. It is submitted that the objections are overcome.

The Office Action objected to claim 71 by asserting that the term “tangible computer-readable medium” is not supported by the specification. By this response, the term “tangible computer-readable medium” is amended to “computer-readable medium.” Appropriate support for the term “computer-readable medium” can be found in, for example, page 11 last paragraph through page 12, first paragraph of the written description.

The Examiner suggested that signal waves are not patentable subject matter. By this response, the specification is amended to clarify that computer-readable medium does not include signal waves.

The Office Action further alleged that “punch cards, paper tape, any other physical medium with patterns of holes,” are mere printed matter and hence are not patentable subject matter. Applicants respectfully traverse. Contrary to the allegation, “punch cards, paper tape, any other physical medium with patterns of holes” are not mere printed matter. Rather, the patterns, holes or any symbols on those media are similar to the pits and lands on a compact disc (CD), which carry meaningful machine-readable instructions (such as logic “0” and “1”) that, upon interactions with and execution by a data processing system, control the data processing system to perform prescribed steps specified by the instructions. As it is well-established that computer-readable medium carrying instructions, such as compact discs, falls comfortably in statutory subject matter,

accordingly, the “punch cards, paper tape, any other physical medium with patterns of holes” also fall in patentable subject matter for the same reasons at least as for the compact discs.

The Rejection under 35 U.S.C. § 112, Second Paragraph Is Overcome

Claim 71 was rejected under 35 U.S.C. 112, second paragraph. Specifically, the Office Action contended that the meaning of the term “tangible computer-readable medium” in claim 71 is not clear. As discussed earlier, the term “tangible computer-readable medium” has been amended to “computer-readable medium” and its meaning is clearly defined in, for instance, the amended paragraph starting from page 11, line 22 to page 12, line 12 of the written description. It is submitted that claim 71 is in proper form.

The Rejection of Claims 53, 54, 56-59, 65 and 71 under 35 U.S.C. § 101 Is Overcome

Claims 53, 54, 56-59, 65 and 71 were rejected under 35 U.S.C. § 101 as allegedly being directed to non-statutory subject matter. By this Response, claims 53, 57, 59 and 61 are amended. It is respectfully submitted that the rejection is overcome.

(1)

The Office Action contended that the following terms are not clear in purpose or scope: “constructing a semantic vector”, “receiving a query containing information”, “comparing the semantic vector for the query to the semantic vector of each dataset”, “selecting datasets whose semantic vectors are closest in distance to the semantic vector for the query”, “outputting information”, “relative strength”, “degree of relevance”, “represents the relative strength”, “using a method other than semantic vectors.”

It is submitted that the following objected terms are removed from the claim language and hence the rejections thereof are moot: “comparing the semantic vector for the query to the semantic vector of each dataset”, “selecting datasets whose semantic vectors are closest in distance to the

semantic vector for the query”, “outputting information”, “relative strength”, “degree of relevance”, “represents the relative strength” and “using a method other than semantic vectors.”

It is further submitted that the meanings of other terms remaining in the claim are clear in light of this amendment and/or the discussions provided below.

a. “constructing a semantic vector”

It is submitted that the term should not be read out of the context of the claims, and that the purpose and scope of the term is clear. According to the claims, it is clear that the purpose of the semantic vectors is to represent datasets or queries. The claims specifically defines how the semantic vectors are constructed by the steps of “for each data point, identifying a relationship between each data point and multiple predetermined categories corresponding to dimensions in the semantic space; defining a formula to calculate signification of each data point with respect to the multiple predetermined categories; determining the significance of each data point with respect to the multiple predetermined categories according to the defined formula; for each data point, constructing a semantic vector representing each data point, wherein each semantic vector has dimensions equal to the number of multiple predetermined categories and represents the significance of its corresponding data point with respect to each of the multiple predetermined categories; and based on the semantic vector for each of the at least one data point, form the semantic vector representing the query or each of the datasets.” The specification also provides ample examples and embodiments on how semantic vectors are constructed and used, which should not be overlooked in interpreting the claims.

b. “receiving a query containing information”

Again, the Office Action improperly read the term “receiving a query containing information” out of context of the claims. The claim language actually reads as “receiving the

query containing information indicative of desired datasets,” and the dataset is defined as “a document.” It is clear from the claim language that a query includes information of documents desired by a user.

(2)

The Office Action alleged that claims 53, 54, 56-59, 65 and 71 because the invention “has not been limited to a substantial practical application,” and selecting abstract datasets or data points is useless in a real world situation.” Applicants respectfully disagree.

The Guidelines for Examination explains that the applicant is in the best position to explain why an invention is believed useful. Accordingly, a complete disclosure should contain some indication of the practical application for the claimed invention, i.e., why the applicant believes the claimed invention is useful. Such a statement will usually explain the purpose of the invention or how the invention may be used (e.g., a compound is believed to be useful in the treatment of a particular disorder).

In determining whether the claim is for a "practical application," the focus is whether the claim, considered as a whole, constitutes "a practical application of an abstract idea." State Street, 149 F.3d at 1373, 47 USPQ2d at 1600. Further, in determining whether the claim is for a "practical application," the focus is not on whether the steps taken to achieve a particular result are useful, tangible and concrete, but rather that the final result is "useful, tangible and concrete." The Federal Circuit further ruled that it is of little relevance whether a claim is directed to a machine or process for the purpose of a Sec. 101 analysis. AT&T, 172 F.3d at 1358, 50 USPQ2d at 1451. Thus, the question of whether a claim encompasses statutory subject matter is "whether the mathematical algorithm was applied in a practical manner." AT&T, 172 F.3d at 1360, 50 USPQ2d at 1453. Accordingly, an "abstract idea" when practically applied to a useful end is eligible for a patent. State

Street, 149 F.3d at 1374, 47 USPQ2d at 1601 ("a process, machine, manufacture, or composition of matter employing a law of nature, natural phenomenon, or abstract idea is patentable subject matter even though a law of nature, natural phenomenon, or abstract idea would not, by itself, be entitled to such protection.") (emphasis added); see also Alappat, 33 F.3d at 1543, 31 USPQ2d at 1556-57(holding that "certain types of mathematical subject matter, standing alone, represent nothing more than abstract ideas until reduced to some type of practical application, and thus that subject matter is not, in and of itself, entitled to patent protection.").

The written description and claims of this application comfortably meet useful, tangible and concrete results. According to the specification, the invention wherein a trainable semantic vector (TSV) is used to provide a semantic representation of information or items, such as documents, in order facilitate operations such as searching, clustering, and classification. See page 4, lines 10-13 of the written description. Techniques described in the application are useable in many real world applications, such as searching documents desired by a user or automatic classifying patents in different classes (see page 31, line 13 through page 43, line 22 of the specification). As described in claims 53, 57, 65 and 71, desired documents are identified and displayed using the unique approach described in the claims. Even assuming that the underlying words or documents are abstract data, the final result of the claims well produces a real-world application allowing a user to retrieve desired documents that are relevant to an input inquiry. It is this end result, not the underlying documents or words, that matters.

Furthermore, the claims of this application produce concrete and tangible results. It is well-established that the tangible requirement does not necessarily mean that a claim must either be tied to a particular machine or apparatus or must operate to change articles or materials to a different state or thing. The claims, as amended, specifically describe that the claims are useful in retrieving

documents that are relevant to a query input by a user. Accordingly, the claims are not directed only to an abstract idea. Rather, they produce real world result by identifying and displaying desired documents to a user. Accordingly, the claims produce tangible results.

It is further submitted that the claims meet “concrete result” requirement. The concrete result requirement only requires that a result to be substantially repeatable. In re Swartz, 232 F.3d 862, 864, 56 USPQ2d 1703, 1704 (Fed. Cir. 2000). The opposite of "concrete" is unrepeatable or unpredictable. The claims clear describe steps and structures that are necessary to implement the inventions and allow any people skilled in the art to repeat the same.

The examiner alleged that the claims are not concrete because “it is not clear what [‘datasets’ and ‘data points’] are,” and “the manipulation of abstract data is not statutory.” See page 6, second paragraph of the office action. To expedite the prosecution process, the claims are amended to clearly identify that a data set is a document and data points are words, phrases or sentences. People skilled in the art can repeat the invention based on the description of the claims and the written description. Whether the underlying words or documents are abstract idea or not is irrelevant to the discussion of producing “concrete results.” As long as the claimed process or system can be repeated or recreated, the claims meet the “concrete result” requirement. Moreover, as discussed earlier, well-established court decisions already pointed out that in determining whether the claim is for a "practical application," the focus is not on whether the steps taken to achieve a particular result are useful, tangible and concrete, but rather that the final result is "useful, tangible and concrete." Thus, it is the result that allows a user to identify documents that are relevant to an input inquiry that matters in producing a useful result. The steps of manipulating data or the underlying words or documents to achieve the intended goal are irrelevant to this inquiry.

The Office Action alleged that the claims are not concrete “because it is not clear what measure of ‘relative strength’ or ‘degree of relevance.’” By this response, these terms are removed from the claims. Furthermore, the claims are amended to clarify that the significance of each data point is calculated by calculating a probability distribution of each data point occurring in each predetermined category and a probability distribution of the data point’s occurrence across all predetermined categories. Appropriate support can be found in, for example, page 16, lines 3-22 of the application. It is submitted that the added description provides a clear and repeatable formula for calculating the significance of the data points. Therefore, the claims produce concrete results.

Additionally, Applicants respectfully submit that claimed process and system, by identifying and displaying documents relevant to an query input by a user, that find relationships between data representing different items in the physical world, are analogous to the transformation of **data** corresponding to discrete dollar amounts into **data** representing a final share price, a case that the Federal Circuit found to be a sufficiently “useful, concrete and tangible result” in *State St. Bank & Trust Co. v. Signature Fin. Group, Inc.*, 149 F.3d 1368, 47 USPQ2d 1596 (Fed. Cir. 1998). Furthermore, processing data related to data points, datasets and queries, and generating representative semantic vectors thereof, to differentiate between related and non-related datasets produces a “useful, non-abstract result” **analogous** to the method of adding **a data field** with information on long distance providers, which the Federal Circuit found to be a “useful, non-abstract result that facilitates *differential* billing of long-distance calls,” which “fall[s] comfortably within the broad scope of patentable subject matter under §101.” *AT&T Corp. v. Excel Communications, Inc.*, 172 F.3d 1352, 50 USPQ2d 1447 (Fed. Cir. 1999).

The Obviousness Rejection of Claims 53, 54, 56-59, 65 and 71 Is Overcome

Claims 53, 54, 56-59, 65 and 71 as being unpatentable over Liddy in view of Wermter. The obviousness rejection is respectfully overcome because Liddy and Wermter cannot support a prima facie case of obviousness.

Embodiments of this application utilize a unique approach to retrieve desired datasets, such as documents, based on a user's search query. A unique type of representation, the semantic vectors, are constructed for the documents and the query. Relevant documents are selected retrieved based on a vector distance between the semantic vectors of the documents and the query. As described in claim 53, the semantic vectors for the query or the datasets (documents) are constructed by the steps of for each data point, identifying a relationship between each data point and multiple predetermined categories corresponding to dimensions in the semantic space, defining a formula to calculate signification of each data point with respect to the multiple predetermined categories, determining the significance of each data point with respect to the multiple predetermined categories according to the defined formula, for each data point, constructing a semantic vector representing each data point, wherein each semantic vector has dimensions equal to the number of multiple predetermined categories and represents the significance of its corresponding data point with respect to each of the multiple predetermined categories, and based on the semantic vector for each of the at least one data point, form the semantic vector representing the query or each of the datasets. The significance of each data point is determined by calculating the percentage of each data point occurring in each predetermined category and the probability distribution of the data point's occurrence across all predetermined categories.

Liddy does not meet the limitations described in claim 53. According to Liddy, each word included in a document looked up in a lexical resource and is assigned a code including a single

category, not multiple predetermined categories. The codes for the words are then combined to form a code for a document. As the types or numbers of words in each document are unpredictable, the generated code for each document in Liddy has variable dimensions, **not multiple predetermined dimensions**, as described in claim 53. In contrast, claim 53 mandates that for each data point (word), a relationship between each data point and **multiple predetermined categories** corresponding to dimensions in the semantic space must be identified, and the identified relationships are utilized to construct a semantic vector having dimensions equal to the number of predetermined categories. Therefore, Liddy fails to disclose “for each data point, identifying a relationship between each data point and multiple predetermined categories corresponding to dimensions in the semantic space;...for each data point, constructing a semantic vector representing each data point, wherein each semantic vector has dimensions equal to the number of predetermined categories and represents the significance of its corresponding data point with respect to each of the predetermined categories; and based on the semantic vector for each of the at least one data point, form the semantic vector representing the query or each of the datasets,” as described in claim 53. Additionally, Liddy does not disclose that the significance of each data point is determined by calculating a probability distribution of each data point occurring in each predetermined category and a probability distribution of the data point’s occurrence across all predetermined categories, as described in claim 53.

The other document, Wermter, cannot be used to modify Liddy because Wermter has a totally opposite goal from claim 53 and hence **teaches away** the techniques described in claim 53. Unlike the semantic vectors described in claim 53, which assigns a document a semantic vector having dimensions equal to **multiple predetermined categories**, Wermter relates to an application called “text routing,” which intends to classify a document into only one **single** category. This is

evident from section 5 of Wermter, in which it describes “in our experiments, we use the recurrent plausibility network described in Figure 1 with two hidden and two context layers. Input to the network is the word representation, one word at a time. Output is the desired semantic routing category.” And in section 6, Wermter states that “the error surface of another title is shown in Figure 4...we can observe that the network has learned the most significant regularities and has a strong preference for a particular semantic category. At the term "Soviet Union", it classifies the title into a wrong category. Then, the word "crude" causes the network preference to switch to the correct category...In the last part of the training, the network has learned the context and gives the correct classification at the end of the title,” and “[i]n Figure 6, we present...This title belongs to the "energy" category...The categorization preference is initially improved but the word "sales" causes an incorrect categorization. However, the words "octane" and "gasoline" cause a switch to the correct category as the correct context is learned.” In other words, Wermter has no intention in creating any multiple-dimensional vector to a document. Rather, the very goal of Wermter is to avoid a situation that a document corresponds to multiple categories, as described in claim 53. Accordingly, Wermter teaches away from meeting the claimed limitations of claim 53, and hence cannot be used to modify Liddy. It is well established that if the proposed modification or combination of the prior art would change the principle of operation of **the prior art**, then the teachings of the references are **not** sufficient to render the claims prima facie obvious. *In re Ratti*, **270 F.2d 810, 123 USPQ 349 (CCPA 1959); MPEP 2143.01(VI)**. Thus, no person of ordinary skill in the art would be motivated by the teachings of Wermter to modify Liddy. Consequently, Liddy and Wermter cannot support a prima facie case of obviousness. It is further submitted that Like Liddy, Wermter also fails to teach that Additionally, Liddy does not disclose that the significance of each data point is determined by calculating a probability distribution of each data

point occurring in each predetermined category and a probability distribution of the data point's occurrence across all predetermined categories, as described in claim 53. The obviousness rejection is untenable and should be withdrawn. Favorable reconsideration of claim 53 is respectfully requested.

Independent claims 57, 65 and 71 also describe assigning a semantic vector having dimension equal to multiple predetermined categories, to a dataset or a data point. Therefore, for at least the same reasons as for claim 53, claims 57, 65 and 71 are patentable over Liddy and Wermter.

Claims 56, 58 and 59 depend on claims 53 and 57, respectively, and incorporate every limitation thereof. Consequently, claims 56, 58 and 59 are patentable over the combination of Liddy and Wermter by virtue of their dependencies. Favorable reconsideration of claims 56, 58 and 59 is respectfully requested.

CONCLUSION

For the reasons given above, Applicant believes that this application is in condition for allowance, and request that the Examiner give the application favorable reconsideration and permit it to issue as a patent. If the Examiner believes that the application can be put in even better condition for allowance, the Examiner is invited to contact Applicant's representatives listed below.

Serial No.: 10/823,685

To the extent necessary, a petition for an extension of time under 37 C.F.R. 1.136 is hereby made. Please charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, to Deposit Account 500417 and please credit any excess fees to such deposit account.

Respectfully submitted,

McDERMOTT WILL & EMERY LLP

/Wei-Chen Nicholas Chen – Reg. No.: 56,665/

Wei-Chen Nicholas Chen

Registration No. 56,665

600 13th Street, N.W.
Washington, DC 20005-3096
Phone: 202.756.8000 WC:MWE
Facsimile: 202.756.8087
Date: November 14, 2007

**Please recognize our Customer No. 20277
as our correspondence address.**

WDC99 1491739-1.055653.0017